California's Landmark AI Law Demands Transparency From Leading AI Developers

What You Need to Know

Key takeaway #1

California recently passed the Transparency in Frontier Artificial Intelligence Act, the first state law to require large AI developers to disclose publicly a safety framework that incorporates widely accepted safety standards and explains a model's capacity to pose, and mitigate, "catastrophic risks."

Key takeaway #2

The TFAIA requires "frontier developers" and "large frontier developers" (both defined in the statute) to make disclosures related to their AI models. The law establishes whistleblower protections for those working for frontier developers. And the law promotes the creation of public infrastructure to support AI research moving forward.

Key takeaway #3

California is the first, but perhaps not the last, state to regulate AI developers specifically. And with the TFAIA, California joins jurisdictions like the European Union, Japan, and South Korea in advancing elements of a risk-based regulatory framework for high-impact AI systems, bringing special regulatory focus to the prospect of catastrophic risk in the U.S. context.

Client Alert | 12 min read | 10.06.25

On September 29, 2025, California Governor Gavin Newsom signed into law Senate Bill 53, the **Transparency in Frontier Artificial Intelligence Act** (TFAIA). This landmark legislation represents California's most significant regulation to date of AI developers.

California is the first state in the country to require large AI developers to disclose publicly a safety framework that incorporates widely accepted safety standards and explains a model's capacity to pose, and mitigate, "catastrophic risks." The law also forces model developers, for the first time, to release transparency reports

on a model's intended uses and restrictions, and it mandates large developers summarize their assessments of a model's catastrophic risks. In addition, the TFAIA breaks new ground by requiring developers to report to the government "critical safety incidents"; by providing whistleblower protections for model developers' employees; and by establishing a consortium to create "CalCompute," a public cloud computing cluster.

Sacramento's move to impose state-level oversight of AI model developers is at odds with **recent** federal actions. That the state with the nation's largest economy and most AI companies is going one direction on AI regulation while the federal government is (**largely**) going another may complicate industry's efforts at compliance.

TFAIA's Narrow But Expanding Applicability

The TFAIA imposes its requirements on a small but growing number of companies. The law applies to "frontier developer[s]"—entities that have trained or are training a "frontier model"—and it applies additional requirements on "large frontier developer[s]," which are frontier developers that have annual gross revenue exceeding $500 million in the prior calendar year. The law defines a "frontier model" by focusing on the amount of computing power that was used to train, fine-tune, or modify the model, namely "a quantity of computing power greater than 10^26 integer or floating-point operations [FLOP]." Sec. 2257.11(i).[1]

Critically, few models today meet the high technical threshold to be a "frontier model," but the trendlines suggest many more will soon. According to **one analysis**, if trends hold, there will be "around 30 such models [that use over 10^26 FLOP] by the start of 2027, and over 200 models by the start of 2030."

The law directs the California Department of Technology to review the definitions of "frontier model," "frontier developer," and "large frontier developer" annually and to submit recommendations to the legislature for updates. Sec. 22757.14(a)-(d).

The TFAIA requires regulated entities to act against "catastrophic risks," which it defines as an incident that causes death or serious injury to more than 50 people or more than $1 billion in damages that involves a frontier model doing any of the following: providing expert-level assistance in the creation or release of weapons of mass destruction; engaging in cyberattacks, murder, or similar crimes; or evading the control of the developer. Sec. 22757.11(c).

Nothing in the law explicitly limits its applicability to frontier developers based in California, and one expects California will attempt to enforce the law on companies that sell their AI products in the state, regardless of their origin, much as California applies its laws to other businesses that have sufficient contacts with the state.

Required Disclosures of Frontier Developers

The TFAIA places meaningful regulatory burdens on frontier developers and large frontier developers by mandating transparency in several ways:

- **Developing and Publishing Frontier AI Framework:** Large frontier developers must write, implement, comply with, and publish on their websites a "frontier AI framework" to manage, assess, and mitigate catastrophic risks. This plan must incorporate "national standards, international standards, and industry-consensus best practices," assess whether the model could pose a catastrophic risk, and mitigate those risks. Among other requirements, the framework must include a description of how the large

frontier developer approaches cybersecurity practices to secure unreleased model weights, which are numerical parameters that determine the strength of connections within a neural network. These disclosures must be updated annually and when "material modification[s]" are made to the framework. Sec. 22757.12(a), (b).

- **Publishing Transparency Report:** Before or when deploying a new frontier model or a substantially modified version of one, a frontier developer must publish on its website a "transparency report" containing information on the model, including its internet address, its release date, its modalities, its intended uses, and "[a]ny generally applicable restrictions or conditions" on the model. A large frontier developer must, in addition, publish a summary of its assessments of the model's catastrophic risks and disclose the extent to which third-party evaluators were involved in those assessments. Sec. 22757.12(c). A large frontier developer must also transmit regularly to the Office of Emergency Services "a summary of any assessment of catastrophic risk resulting from internal use of its models." Sec. 22757.12(d).

- **Incident Reporting:** Frontier and large frontier developers must within fifteen days of discovering a "critical safety incident" report that incident to the California Office of Emergency Services. Critical safety incidents include the unauthorized access, modification, or exfiltration of the model weights of a frontier model that results in death or bodily injury; harm resulting from the materialization of a "catastrophic risk"; the "loss of control" of a frontier model causing death or injury; and, when a frontier model uses deceptive techniques to "subvert the controls or monitoring" of the model "in a manner that demonstrates materially increased catastrophic risks." Secs. 22757.13(c), 22757.11(d). The law encourages but does not require frontier developers to report critical safety incidents pertaining to foundation models "that are not frontier models." Secs. 22757.13(c)(4). The Office of Emergency Services will also establish a mechanism for the public to report such incidents. Sec. 22757.13(1). Beginning January 1, 2027, the Office of Emergency Services will produce an annual report of anonymized and aggregated information about critical safety incidents. Secs. 22757.13(g).

- **Deemed Compliance:** The law contains a provision that will allow companies to comply with the incident-reporting provisions circuitously. If the Office of Emergency Services designates a *federal* law, regulation, or policy as imposing standards equivalent to, or stricter than, the TFAIA's incident reporting standards, and the frontier model complies with the federal standard, the frontier developer "shall be deemed in compliance" with the California state law, even if the federal policy does not expressly preempt state law. Sec. 22757.13(h), (i).

- **Penalties:** Large frontier developers that make false statements, fail to tender required reports, fail to report critical incidents, or fail to comply with their own frontier AI frameworks could be subject to a civil action brought by the California Attorney General with civil penalties of up to $1 million per violation. Sec. 22757.15(a), (b).

Whistleblower Protections

The TFAIA establishes whistleblower protections for those working for frontier developers:

- **Protected Disclosures:** The TFAIA establishes protections for employees, contractors, and affiliates of all frontier developers who report violations by developers of the law or raise concerns about safety

and ethical risks associated with frontier models. Sec. 1107.1(a)-(e).

- **Anti-Retaliation:** The law prohibits employers from retaliating against individuals who make protected disclosures. If retaliated against, whistleblowers may seek legal remedies. Sec. 1107.1(g).
- **Confidentiality:** The law requires large frontier developers to provide internal processes for whistleblowers to report information anonymously. Such disclosures will be shared with officers and directors of the large frontier developer, subject to certain exceptions if the whistleblower alleges a director or officer of wrongdoing. Sec. 1107.1(e).
- **Attorneys' Fees:** Plaintiffs that show that whistleblower protections have been violated may be entitled to attorneys' fees. Sec. 1107.1(f).

Public Cloud Computing Cluster

In addition to regulating the development of frontier models, the law promotes the creation of public infrastructure to support AI research moving forward:

- **Consortium:** The law establishes a public-private consortium to develop recommendations for CalCompute, a state-managed, open-access cloud computing cluster. Sec. 11546.8(a).
- **Purpose and Goals:** CalCompute is intended to provide secure, equitable, and sustainable computing resources for AI research and development, with a focus on advancing projects that benefit the public and meet high ethical standards. Sec. 11546.8(b).
- **Stakeholder Engagement:** The consortium will include representatives from academia, industry, government, and civil society, ensuring broad input into the design and governance of CalCompute. Sec. 11546.8(g).

California May Serve as a Model for Other States and Jurisdictions

California is the first, but perhaps not the last, state to regulate AI developers specifically.

In New York, a bill awaiting Governor Kathy Hochul's signature, the **Responsible AI Safety and Education (RAISE) Act**, would require developers of frontier AI models to create and maintain safety and security protocols, report significant safety incidents to the state, evaluate their models and withhold any model that poses an "unreasonable risk of critical harm," such as mass casualties or significant economic damage, and enlist a third-party to perform a yearly, independent audit of the developer's compliance with the law.

Other states have enacted AI laws that regulate AI when it is used in particular ways. (This approach differs from the TFAIA, which regulates AI based on whether the AI model itself is advanced, or "frontier.") For example, the **Colorado Artificial Intelligence Act**, which will go into effect on June 30, 2026, imposes various obligations related to documentation, disclosures, and governance of "high-risk" AI systems—systems that make "consequential decisions" relating to education, employment, health care, and similar areas. The **Utah Artificial Intelligence Policy Act** took effect on May 1, 2024, and was **amended** recently to require certain disclosures for "high-risk" AI interactions. In May 2025, Texas enacted the **Texas Responsible Artificial Intelligence Governance Act**, which prohibits the development and deployment of AI that engages in certain conduct, namely intentionally inciting self-harm or criminal activity or violating users' Constitutional rights.

Many states have also adopted sector-specific AI-laws that regulate the use of AI in particular domains, such as, **chatbots**, **insurance**, **nonconsensual intimate imagery**, **political advertisements**, and **healthcare**.

Several additional laws in California will also go into effect on January 1, 2026, including: the **AI Training Data Transparency** (AB 2013), which requires AI developers and those who "substantially modif[y]" or tune AI models to post publicly details on the data on which their models were trained; the **California AI Transparency Act** (SB 942), which requires businesses providing generative AI systems with over a million monthly visitors to provide a free AI detection tool for users and include disclosures on AI-generated media; [2] and **Automated Decision-Making Technology Regulations**, issued by the California Privacy Protection agency, which will require businesses to inform workers when automated technology is used to make decisions affecting employment, conduct risk assessments, and offer consumers and customers a way to opt-out of their use.

With respect to international benchmarking, as a major economy itself, California joins jurisdictions like the European Union, Japan, and South Korea in advancing elements of a risk-based regulatory framework for high-impact AI systems, bringing special regulatory focus to the prospect of catastrophic risk in the U.S. context. The passage of the TFAIA also comes on the heels of the United Nations' launch of the **Global Dialogue on Artificial Intelligence Governance**, announced in late September, which emphasizes alignment and cooperation on policy, science, and capacity building on AI within UN circles. The novel lens the TFAIA applies to frontier technologies and the distinct requirements it imposes on a narrow subset of developers independent of risk assessments will no doubt become the subject of debate in state and international regulatory dialogues to come.

The TFAIA represents a watered-down version of an AI safety bill that Gov. Newsom vetoed in 2024 (**SB 1047**), which would have required third-party audits of frontier models, testing and certification prior to a model's release, and the inclusion of a "kill-switch" that would shutdown models. The TFAIA's narrow focus on transparency and whistleblower protections may serve as a model to other states that are rushing to pass AI-related laws. The federal government considered but dropped a so-called "**moratorium**" on state AI laws earlier this year. The White House's **AI Action Plan**, published in July 2025, has sought to review federal AI regulations that are "burdensome" to AI development, but major legislation either to ease regulatory burdens on AI or preempt state laws remains speculative.

In a **statement** accompanying his signature of the TFAIA, Gov. Newsom wrote that, should the Federal government or Congress adopt national AI standards that meet or exceed the TFAIA's protections, "subsequent action will be necessary to provide alignment between policy frameworks[.]"

Next Steps

Frontier developers and large frontier developers should review the TFAIA and consult with counsel to assess the law's impact on their operations. These entities may seek to assess their current transparency practices; draft frontier AI frameworks and transparency reports, where appropriate; formalize processes to report critical safety incidents; assess catastrophic risks and enhance internal documentation; update whistleblower policies and notify workers of whistleblower protections; track California's designation of compliant federal, incident-reporting policies, if any; and monitor developments related to CalCompute.

For more on California's AI laws and national trends in AI regulation, please join Crowell & Moring for a webinar on October 16, 2025 on "**The Artificial Intelligence Agenda from Capitol Hill to State Capitals: Where We Are and Where We Are (Probably) Going**."

[**1**] This was the same computing threshold for compliance reporting required by the **Executive Order on AI** that President Biden signed in October 2023. It exceeds the computing threshold of the **EU AI Act**, which imposes regulations on general purpose AI models that are trained on cumulative compute greater than 10^25 FLOP.

[**2**] In mid-September 2025, the California legislature passed a **bill** (AB 853) that, if Governor Newsom signs it into law, will delay the operation of the California AI Transparency Act until August 2, 2026.

Contacts

Matthew F. Ferraro
Partner
Washington, D.C. D | +1.202.624.2610
mferraro@crowell.com

Justin B. Weiss
Crowell Global Advisors Senior Director, Senior Counsel
Washington, D.C. (CGA) D | +1.202.654.6729
jbweiss@crowellglobaladvisors.com

Matthew Moisan
Partner
He/Him/His
New York D | +1.212.223.4008
mmoisan@crowell.com

Meaghan Katz
Associate
They/Them/Theirs
San Francisco D | +1.415.365.7287
mkatz@crowell.com

Jacob Canter
Counsel
He/Him/His
San Francisco D | +1.415.365.7210
jcanter@crowell.com

